DOCUMENT RESUME

ED 229 768 CS 207 500

AUTHOR Takala, Sauli
TITLE Student Writing: Some Notes on Definition and Measurement.
PUB DATE Mar 83
NOTE 25p.; Paper presented at the Annual Meeting of the Conference on College Composition and Communication (34th, Detroit, MI, March 17-19, 1983). Some charts may not reproduce well due to small print.
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cross Cultural Studies; Elementary Secondary Education; *Evaluation Criteria; Foreign Countries; Holistic Evaluation; *Linguistic Competence; *Measurement Techniques; Models; Student Behavior; Writing (Composition); *Writing Evaluation; *Writing Research

ABSTRACT

The first task of the International Association for the Evaluation of Educational Achievement's International Study of Written Composition was to conceptualize the writing domain by determining the most significant parameters that have to be taken into account in all writing situations. A model was developed that used the level of cognitive processing as one dimension and the purposes of writing as the other. The dependent measure, "student writing activities," was divided into two related parts: writing competence and writing preference. Three populations were tested and subjects were asked to participate in nine writing tasks: pragmatic information, summary, retelling a story, descriptive composition, personal story, argumentative/persuasive, personal/reflective, open composition to pictorial stimuli, and a letter of advice. Compositions were scored in terms of general and detailed impressions. A small-scale pilot study in which native language experts from nine countries rated argumentative and reflective compositions showed a high agreement on all the rating categories (based on competence in generating and organizing ideas; applying appropriate style and tone; using appropriate grammatical, lexical, and spelling conventions; and producing a legible text). This, together with the fact that the raters were able to communicate with relative ease using the categories, suggested that the categories of writing competence were familiar and appropriate in various parts of the world. Writing preference, determined by self-reporting techniques and descriptive analysis of students' compositions, was also found to be a variable worthy of consideration in developing a system for assessing student writing activities. (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED229768

STUDENT WRITING: SOME NOTES ON DEFINITION AND MEASUREMENT

Sauli Takala

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sauli Takala

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Curriculum Laboratory
University of Illinois at
Urbana-Champaign

Paper presented at the Annual Meeting of the Conference on College Composition and Communication, Detroit, March 17-19, 1983

CS 207 500

Contents

		Page
1.	Introduction	1
2.	Towards the Conceptualization of the Domain of Writing	2
3.	Towards the Conceptualization of Student Writing	5
	3.1 Writing Competence	5
	3.2 Writing Preference	7
4.	Selection and Specification of Writing Tasks	10
5.	General Principles of Scoring the Compositions	13
6.	Scoring of Compositions Across Cultures and Languages: Some Pilot Findings	14
7.	Conclusion	21
	References	23

1. Introduction

Under the auspices of the International Association for the Evaluation of Educational Achievement (IEA) empirical cross-national studies of educational achievement have been carried out during the past twenty years or so. About three years ago the General Assembly of the IEA approved of a proposal to initiate an international study of written composition.

It soon became evident that the project managers needed to do a lot of conceptual exploration and clarification in order to be able to carry out its main objectives: to describe the current state of instruction in written composition in various parts of the world, to assess student performance in written composition, and to explain what factors account for differences in student performance. One of the most difficult and central problems facing the project has been the definition and measurement of the dependent variable. On the surface, it seems that the dependent variable is fairly obvious: ability to write compositions. A closer look reveals, however, that the nature of the dependent variable is not so self-evident but requires careful conceptual analysis.

This paper seeks to contribute to such a conceptual clarification. It also seeks to show how the selection and specification of the tasks and the scoring of the compositions are related to the conceptualization of the dependent variable.

2. Towards the Conceptualization of the Domain of Writing

As in all testing and evaluation, a central demand for the IEA International Study of Written Composition is that it is adequate in terms of construct and content validity. The writing tasks have to take into account the nature of writing as a cognitive and social act. This is necessary to establish a solid theoretical foundation for the study (construct validity). A theoretical model of the domain of writing is needed to guide the principled selection of tasks from the universe of all possible writing tasks. This overall framework also helps in overcoming the serious problems of curricular validity in a situation where a set of tasks is needed that is a fair sample of typical writing tasks in a number of different countries. Without such a theoretical model that helps in creating a common framework (schema) for discussions and in establishing a common set of terms, it would be extremely difficult to arrive at a set of tasks that can be justified both theoretically and in terms of curricular validity. To go the other, empirical, way by studying the curricula and typical writing assignments, examination topics, etc. would have been extremely laborious and would have necessitated the construction of some taxonomic classification system anyway.

The starting point of the IEA Study of Written Composition is distinctly theoretical. It was assumed that in a poorly defined area like the teaching of written composition the most useful thing to do first was to try to conceptualize the domain: to define the domain of writing by determining the most significant parameters that have to be taken into account in all writing situations. If such a conceptual model is theoretically sound, it should prove a powerful aid in communicating across cultures, languages, and national school systems. It should help in selecting tasks and defining their specifications. It should also help in securing a relatively high degree of interrater agreement across different school systems.

After several drafts, which emphasized the purposes and contents of writing, it was decided that a more satisfactory model would be one that uses the level of cognitive processing as one dimension and the purposes of writing as the other. The first dimension proceeds from low to high demands on cognitive processing. Thus, in reproduction the student is given all

he or she needs to produce the final product. Neither the generation (invention) of content nor of the linguistic expression is required. In the following stage, organization or reorganization, the student knows or is given the content, is given cues as regards format etc., but has to transform both the content and the language in order to produce an appropriate product. The most demanding level of cognitive processing, generating or inventing, requires that the student generates both the content and the linguistic expressions on the basis of his or her own mental resources. The model (Vähäpassi 1982) is shown in Chart 1. It has to be noted that it is a general model and that the line between organizing/reorganizing and inventing/generating is conceptual. A particular task may be one of organizing or inventing depending on the cognitive history of each individual.

The model differs from most of the previous in that the two dimensions also have related sub-dimensions. Both main dimensions consist of verbs, in most cases clearly transitive verbs with specifiable objects. The cognitive processing verbs are related with three broad classes of content: time-related events and processes, space-related objects, and (for lack of a better term) mentally related ideas, mental structures and systems. We manipulate such content in thinking and in writing. The intentional action verbs, which indicate the dominant purpose of writing, are closely related to whom the intentional action is aimed at: mainly self or others.

One of the merits of this kind of taxonomy is that it forces us to pay attention both to the intentional character of action and the cognitive processing involved in intentional action. Thus it helps to understand why it is questionable to say that much of school-based writing is writing without a "real" purpose or audience. Typical writing done in worksheets or in examinations can be described in terms of the model as follows: reproduce facts/ideas in order to inform the teacher about the progress of learning. Such writing provides useful diagnostic information for the teacher and for the student. Much school writing can also be described in the following way: reorganize facts/ideas in order to convince the teacher that the student deserves a certain grade. Some kinds of writing assignments might be characterized thus: generate ideas/hypotheses and work them out in order to learn (gain new insights). In this case the primary audience of writing is the

Chart 1. Tasks in Relation to the Domain of Writing (slightly modified from Vähäpassi 1982)

Cognitive Processing / Dominant Intention/ Purpose	Primary Audience	REPRODUCE	ORGANIZE/REORGANIZE		INVENT/GENERATE	
	Pri- Con- mary tent	Facts Ideas	Events	Visual images mental states, facts, ideas	Ideas, mental states, alternative worlds	
To learn (meta-lingual / interpretive)	S e l f	Copying Taking Dictation	Retell a story (heard or read) (2)	Worksheets Note Resume Summary (3) Outline Paraphrasing	Comments on book margins Metaphors Work out an idea/ Analogies hypothesis Reflective essay (7)	
To convey/ signal (emotive)	Self Others	Stream of consciousness	Personal story (5) Personal diary Personal letter (8)	Portrayal/ characterization	Reflective writing --Personal essays (7)	The traditional literary genres and modes can be placed under one or more of these four purposes.
To inform (referential)	O t h e r s	Fill in a form Use quotation Exam answers Message (1)	Exam answers Narrative report News Instruction Telegram Announcement Circular Message (1)	Exam answers Directions Description (4) Technical description (9) Biography Science report/ experiment	Expository writing --Definition --Academic essay/ article (7) --Book review --Commentary	
To convince/ persuade (conative)	O t h e r s	Citation from authority/expert	Exam answers Letter of application (resume/vita) Statement of personal views, opinions Exam answers (8)	Exam answers Advertisement Letter of advice (9)	Argumentative/ persuasive writing (6) --Editorial --Critical essay/ article (7)	
To entertain, delight, please (poetic)	Others	Quote poetry and prose	Given an ending, create a story Create an ending Retell a story (5)	Word portrait or sketch Causerie (8)	Entertainment writing --Parody --Rhymes --Story (8)	
PRIMARY MODE OF DISCOURSE		DOCUMENTATIVE DISCOURSE	CONSTATIVE DISCOURSE Narrative Descriptive Explanatory		EXPLORATORY DISCOURSE Interpretive Literary (Expository/ Argumentative)	

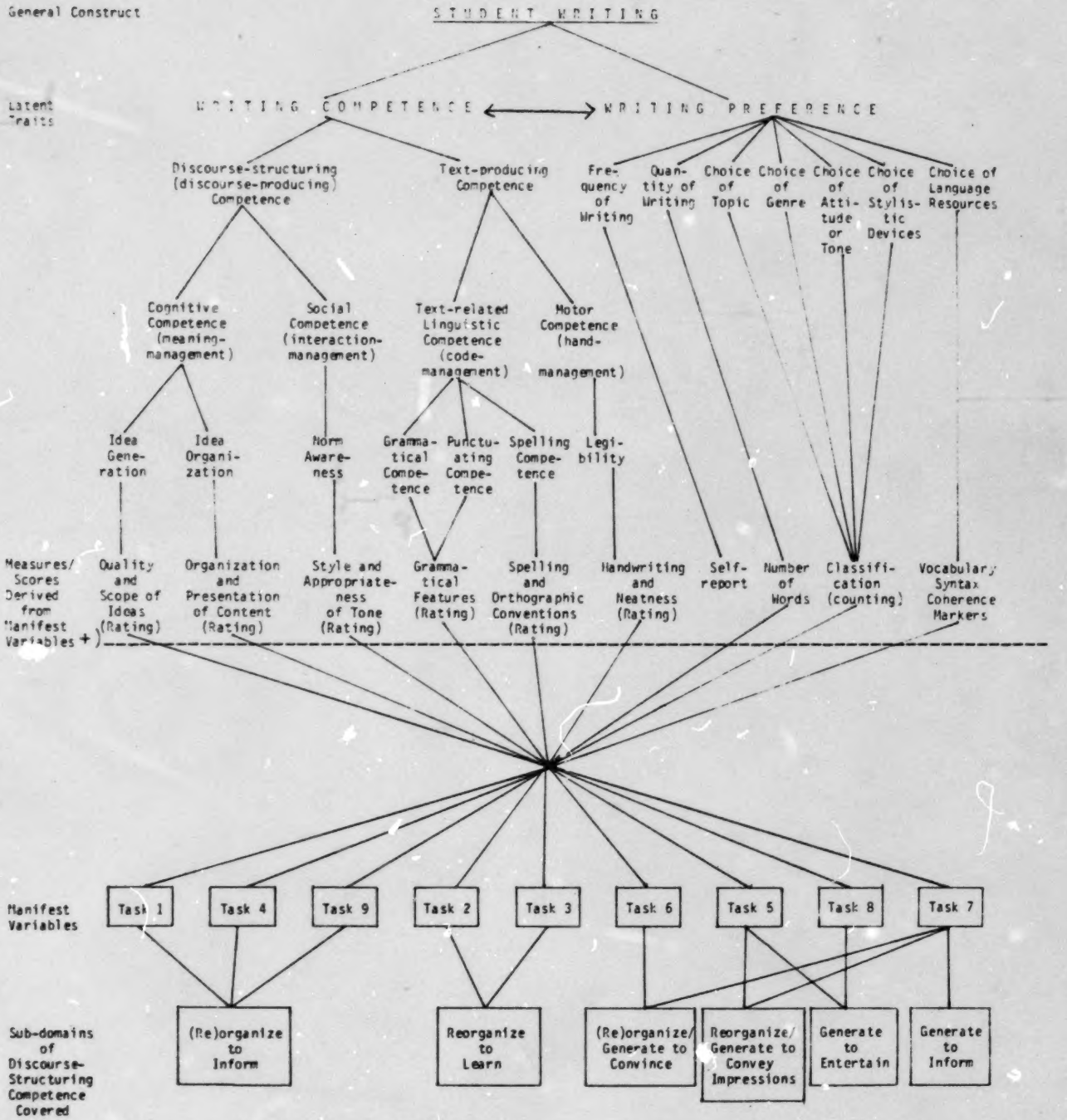

Figure 1. Structure of the Dependent Variable: Student Writing

+) An overall impression score as well as a rating of reader's response (eg. interest, sense of being persuaded etc), which probably are based on several or all of these sub-categories (and probably with somewhat different weights for different readers) could be added here as measures of rated writing performance.

student himself but the teacher can again give comments to help the student to see the weak and strong points in such an expository use of writing. An educational institution should not underestimate the use of writing as a tool of learning. Nor should it underestimate the student and his or her teacher as the audience of writing. There is nothing "unreal" about students, their teachers and their schools. On the contrary, the schools are real places, which are doing a creditable job in enhancing students' mental growth by providing a systematic and structured program of activities, which the "real" out-of-school "program" with its more random or more narrow focus can hardly match.

3. Towards the Conceptualization of Student Writing

The dependent measure of the IEA Study of Written Composition is a broad concept, which might be called "student writing activities" or perhaps better "student writing". It is divided into two related parts, which can, however, be held to be conceptually distinguishable: "writing competence" and "writing preference". The structure of the dependent variable is illustrated in Figure 1.

3.1. Writing Competence

"Writing Competence" can be operationalized as the ability to produce written products that cover the cells of the typology of the writing domain (Chart 1). A person may be able to write fluently a given type of discourse (e.g. a story, a personal letter, an academic paper) because he or she has a well-developed schema for that. Such a person may thus be called a competent or fluent story-writer or letter-writer but not a competent writer, if his or her competence is limited. A competent writer needs to be able to write across a large range of tasks and thus be flexible (Purves 1981) and be able to use the linguistic resources with required correctness and appropriateness and exhibiting sufficient fluency.

Writing competence, in turn, can be argued to consist of two main components: discourse-structuring competence (or discourse-producing competence) and text-producing competence. Discourse-structuring competence refers to the cognitive ability to encode meanings and intentions effectively.

Discourse-structuring competence means the ability to generate written products in which the units of thought and the units of language are related to each other in such a way that an appropriate structure of meaning is produced. The appropriateness is always dependent on the intention of the writer and the audience of writing: appropriateness is not a universal concept, it is always context and situation specific.

It is important that the ideas are perceptive, relevant and clear for the audience of writing (idea generation) but they must also be arranged in such a way that a discourse type is recognized (idea organization). It is not irrelevant how the meaning (deep structure) is organized in a linear text (surface structure). Ease of comprehension and recall are usually better if the two coincide. Brewer (1982) has shown, however, that the events in a story have to be arranged in a certain order for the story to produce either suspense, surprise or curiosity in readers. Readers have genre structural knowledge and expect sufficient conformity with typical genre schema. Discourse has to be structured differently if the purpose (intention) changes from narrative to persuasion or to description.

Since writing is usually addressed to an audience other than self, writing competence (discourse-structuring competence) also presupposes social competence. The writer has to be aware of audience expectations (norms) and use an appropriate tone and style.

Text-producing competence can be divided into two parts: linguistic competence and motor competence. Linguistic competence consists of the ability to use appropriate grammar and spell and punctuate properly. Motor competence refers to the ability to produce an easily legible text.

3.2 Writing Preferences

Another aspect of student writing, which is complementary to students' writing competence, is their writing preferences. The former is normative in nature and can be rated in accordance with certain criteria, on which a relatively high agreement is possible to obtain after discussion and training. All written products can be judged in terms of their competence or at least their merits and weaknesses can be critically discussed. If competence is indicated simply by assigning one score or a set of sub-scores, the score or scores usually tell us only the raters' quality judgments. They do not typically

describe what there was in the composition that led to the quality ratings. Thus several compositions may have identical ratings but they may still be quite different.

It is possible to make a list of characteristics that are related to various aspects of compositions and indicate the presence or absence of the frequency of the traits. Such a checklist is basically descriptive in nature. To indicate the absence or presence of a trait presupposes the ability to identify and classify traits in a reliable manner. Some traits can be indexed through simple counts but others may require the ability to make fine distinctions.

In the case of a type of writing which is strongly governed by a set of conventions, a detailed description of a piece of writing may lead to a conclusion about its merit also. In other types of writing there may exist a correlation between description and quality rating but the relationship is less determinate. Although description and judgment are to some extent related and description may imply judgment, it is useful to keep them conceptually apart because doing so provides additional information. Judging is like summative evaluation, description is more like diagnostic data, which can be used for evaluative purposes. In the domain of human action probably any descriptive data can be, and a lot of them often are, used for quality judgments.

Writing preference, as any preference, can reliably be gauged only if there are several occasions on which behavior is observed. One instance of behavior may indicate preference but it may also indicate exceptional behavior, a deviation from the typical. If we wish to get an account of individual writing preferences we have to get several samples of his or her writing or to get a self-report on what he or she usually chooses to do. We may also get an idea of what the preferences are in a given speech community (or rhetorical community) by observing a representative sample of individuals of that community.

One aspect of writing preference is the topic that students choose to write on when they have a choice. What are the areas which they deal with in their writing and what aspects they select for discussion. The topics and sub-topics can be classified to provide both measures of central tendency and measures of dispersion. Such an analysis would show, for example, whether

students concentrate on a few points or whether they prefer to deal with several aspects. Information on this sub-area of writing preference can be obtained both by self-reporting and by analyzing a number of students' compositions. The students' choice of genre or mode is also an indicator of writing preference. If they have a choice, do they prefer an imaginative or a more matter-of-fact mode of writing? Do they prefer a narrative mode to a more expository one?

The overall tone of writing is another aspect of writing preference. Do students indicate a definite stand in an assertive way (positive, negative) or are they more detached, more impersonal, more non-committed (presenting pros and cons but not going beyond that)? Do students use humor, wit, irony, satirical remarks or is their writing mostly serious?

The above are related to matters that are to do more with non-language factors (personality, interests, etc.) than with the choice of linguistic options. The latter can be divided at least into two groups: the choice of rhetorical devices and the choice of language resources.

The choice of rhetorical devices refers to the extent to which students use metaphors, similes, etc. How often do they quote aphorisms, proverbs, lines from poetry, etc? To what extent do they use repetition for stylistic effect?

The choice of language resources refers to the kind and amount of vocabulary they typically use, the type of sentence patterns they prefer, and the ways they use language to make their text cohesive and coherent.

We can study preferences in all circumstances. Even when the writing tasks are defined in great detail, a student still has a set of options to choose from. When students write on their own or when the assignments given to them allow a wide choice of approach and the criteria of good performance are not strictly imposed on students, we can assume that what they write in such circumstances reflects their natural writing preferences, or preferences that they have learned to be successful.

Preference by definition implies that there must be a choice: behavior is not completely determined by the circumstances. Writing preference has been conceptually defined in Figure 1. All sub-areas that constitute the concept "writing preference" can be described by numerical indices. The most prominent sub-areas of writing preference are briefly discussed in the following (for a more detailed account, see Takala 1982).

When students have a choice of writing or doing something else, the frequency of their writing activities is a clear indicator of their writing preference. Information on the frequency of writing can be obtained by observation and by self-reporting. Voluntary writing can be treated both as an independent and dependent variable. As an independent variable, together with non-voluntary frequency of writing, it can be assumed to influence writing competence. Thus it can be used as a predictor in regression and path analysis. It is also conceptually possible to regard voluntary frequency of writing as, at least partly, the outcome of writing instruction. Since writing is typically learned at school, the school can presumably take partial credit for voluntary out-of-school writing and partial blame for avoidance of writing. This, of course, presupposes that writing is considered a desirable activity. Arguments to support this view have been presented among others by Emig (1977), Bruner (1972), and critically discussed by Scardamalia and Bereiter (1981).

The quantity of writing, also a matter of choice, can be assessed most simply by counting the number of words. This poses, however, the problem of what is counted as a word. In some languages, which attach suffixes to the end of the words, the number of words is lower than in those languages which use e.g. articles and prepositions, if a word is counted mechanically as a string of letters which are separated by space. Another way of quantifying the amount of writing would be to identify the number of prepositions.

4. Selection and Specification of Writing Tasks

One of the greatest weaknesses of evaluation studies has been that the domain of the knowledge and skills to be tested has not been carefully defined. In large-scale evaluation studies the representativeness of the student sample has been an object of conscious attention, since the goal has been to be able to generalize the level of achievement to the whole population and to its various significant strata. Thus it has been possible to say that some group is doing better than another group. Only relatively recently has the representativeness of the test content been taken equally seriously and the sampling from the item universe has assumed an equally important

position. Hence, the selection and specification of writing tasks was a central concern in the IEA Study of Written Composition.

In order to be able to estimate development in writing and in order to test some hupotheses three populations are tested (some countries test only one popula;tion and some only two). Population A is defined as students at or near the end of primary education and the self-contained classroom. Population B is defined as students at or near the end of comprehensive education, i.e. students who are in the last year of shortes secondary program and those in longer programs who have completed the same number of years of schooling whether or not they have finished their program. Population C consists of those students who are at or near the end of academic secondary school (pre-university year).

The allocation and linkage of tasks among the three populations is as follows:

		Populations
Task 1:	Pragmatic Information	
	T1A: Description of a Bicycle	A & B
	T1B: Self-description	A & B
	T1C: Letter to a Principal	A & B
	T1D: Note to Family	A
	T1E: Application for a Holiday Job	B & C
Task 2:	Summary	B & C
Task 3:	Retelling a Story	A
Task 4:	Descriptive Composition	
	T4A: Description of a Mask	A & B
	T4B: Description of a Process	A & B
Task 5:	Personal Story	A & B
Task 6:	Argumentative/persuasive Composition	A & B & C
Task 7:	Personal/reflective Essay	B & C
Task 8:	"Open Composition" (pictorial stimulus)	B & C
Task 9:	Letter of Advice	B & C

The tasks are briefly discussed below indicating also how they are related to the domain of writing (Chart 1).

Task 1 - Pragmatic Information

Students are asked to write a short practical message to some clearly identified addressee in a realistic communicative situation. The primary purpose is referential: the purpose is to inform the addressee. The cognitive demand is to reproduce or (re)organize facts that are given to the students. This task reflects a pragmatic or functional, minimum-competency type of writing.

Task 8 - "Open Composition"

The students are asked to respond with maximum personal freedom to a pictorial stimulus with no clear referential "message." Thus the task is open for students to interpret in any way they like. The cognitive demand is to invent the ideas and to generate the linguistic expressions to suit the ideas. The purpose of writing may vary from conveying emotional states to entertaining oneself or the reader.

Task 9 - Letter of Advice

The students are asked to write a letter to a younger student who is assumed to be coming to their school advising him or her how to write compositions in their school so as to get good grades. The students are "experts" in this area so the cognitive demand is to reorganize facts and ideas. The dominant purpose is to inform the younger student.

A look at the tasks and Figure 1 shows that the tasks cover a wide range of the domain of writing. The referential function of language, when the writer tries to inform the reader, is, however more prominent than the other language functions. This is in accordance with the curricular emphasis in the participating countries. In several cases, particularly Task 8 but also to some extent Tasks 5 and 7, the task instructions are purposefully relatively non-specific in order to allow students a wider choice in the way they wish to treat the assignments.

5. General Principles of Scoring the Compositions

After lengthy discussions and exploratory work related to scoring compositions from several countries and written in several languages it was decided to score all compositions both in terms of a general impression and in terms of more detailed impressions. Thus the scoring system combines both a holistic and analytic scoring principles. The main reasons for this decision were that using the overall impression provides a common metric across tasks and it also reflects possible national assessment preferences. The more analytic scoring, on the other hand, provides more detailed - diagnositc-evaluative - information about some central aspects of compositions.

The scoring systems are essentially the same across all tasks. The only difference is that some tasks are scored in greater detail than others. It was first considered that the pragmatic information tasks, the various versions of Task 1, might be scored by just noting whether some crucial points are present in the composition or not (e.g. date, time, place of meeting). Discussions showed, however, that the same criteria of the adequacy of content, organization and style/tone apply to them as well as to the less pragmatic/ transactional type of writing.

These considerations, some empirical pilot work and earlier studies on the scoring of compositions led to the six categories shown in Figure 1: Quality and Scope of Ideas, Organization and Presentation of Content, Style and Appropriateness of Tone, Grammatical Features, Spelling and Orthographic Conventions, Handwriting and Neatness. A seventh category, Response of Rater, was added to let the rater indicate his or her interest in the composition, whether the writer senses a personal voice in the composition or - as in Task 6 - whether the writer has succeeded in persuading the rater that he or she has made a good case for his or her opinions.

In Tasks, 1, 2, 3, 4 and 9 there is no further elaboration of the seven rating categories whereas in Tasks 5, 6, 7 and 8 the raters are asked to give first a rating on each of the seven categories and immediately after that attempt to provide a more detailed rating within each category (see Scoring Sheets on pages 17 and 18).

6. Scoring of Compositions Across Cultures and Languages: Some Pilot Findings

After several days of discussions dealing with the development of the scoring systems for the various tasks and after a joint selection of sample essays to illustrate the various scale point values (from low to high on each category) accompanied by discussion of reasons for assigning such values to the chosen compositions, a group of experts of mother tongue teaching from nine countries proceeded to rate compositions. Fifteen compositions on Task 6 were rated and ten compositions on Task 7. The compositions were written by students in a number of countries and some of them were translated into English. Since all the compositions were typed and since some of them were

Sample Scoring Sheet
[Tasks 1, 2, 3, 4, 9]

IEA INTERNATIONAL STUDY OF WRITTEN COMPOSITION

Population_____ Task_____ Student_____ Rater_____

WITH RESPECT TO WRITER'S CHOICE OF AIM AND READERSHIP

		Inadequate				Excellent
		1	2	3	4	5
A.	Quality and Scope of Content	___	___	___	___	___
B.	Organization and Presentation of Content	___	___	___	___	___
C.	Style and Appropriateness	___	___	___	___	___
D.	Grammatical Features	___	___	___	___	___
E.	Spelling and Orthographic Conventions	___	___	___	___	___
F.	Handwriting and Neatness	___	___	___	___	___

		Low	Medium	High
G.	Response of Rater My interest in this composition is	___	___	___

Sample Scoring Sheet

TASK 6 - ARGUMENTATIVE/PERSUASIVE COMPOSITION

Population _____ Student _____ Rater _____

WITH RESPECT TO WRITER's CHOICE OF AIM AND READERSHIP

		Inadequate 1	2	3	4	Excellent 5
A.	Quality and Scope of Content					
	1. Significance of what is said					
	2. Argumentation/Exposition					
B.	Organization and Presentation of Content					
	3. Organization of the whole text					
	4. Organization of sub-units					
C.	Style and Tone					
	5. Choice of consistency of tone					
	6. Choice of words and phrases, sentence structures, and larger units of discourse					
D.	Grammatical features					
E.	Spelling and orthographic conventions					
F.	Handwriting and neatness					

G. Response of Rater	Low	Medium	High
8. My interest in the composition is			
9. My sense of being persuaded by the composition			

TABLE 1. Reliability coefficients (alpha) of different rating categories (Task 6, 9 raters and 15 compositions)

OVERALL IMPRESSION	.944
QUALITY AND SCOPE OF CONTENT	.926
Significance of problems and solution	.910
Success of argumentation, persuasiveness	.934
ORGANIZATION AND PRESENTATION OF CONTENT	.936
Degree to which thesis is reflected in the whole text	.934
Consistency and control of supporting arguments	.939
STYLE AND APPROPRIATENESS	.951
Choice of words and sentences	.959
Choice of consistency of "tone"	.943
RESPONSE OF RATER	.905
My interest in the composition	.916
My sense of being persuaded by the composition	.894
TOTAL	.926

TABLE 2. Reliability coefficient (alpha) of different rating categories (Task 7, 9 raters and 10 compositions)

OVERALL IMPRESSION	.928
QUALITY AND SCOPE OF IDEAS	.926
Significance of what is said	.899
Penetration in exploration of ideas, depth	.943
ORGANIZATION AND PRESENTATION OF CONTENT	.929
Use and placement of detail	.935
Organization of the whole text including cohesion	.926
STYLE AND APPROPRIATENESS	.903
Choice of words and sentences	.914
Choice and consistency of "tone"	.905
RESPONSE OF RATER	.885
My interest in the composition	.902
My sense of connection with the writer as a person	.869
TOTAL	.905

TABLE 3. Intercorrelations among various rating categories. (Task 6, 9 raters and 15 compositions)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. OVERALL IMPRESSION	1.00												
2. QUALITY AND SCOPE OF CONTENT	.88	1.00											
3. Significance of what is said	.82	.95	1.00										
4. Argumentation/exposition	.86	.96	.81	1.00									
5. ORGANIZATION AND PRESENTATION OF CONTENT	.88	.88	.79	.87	1.00								
6. Organization of the whole text	.85	.86	.79	.84	.97	1.00							
7. Organization of subunits	.85	.84	.74	.86	.97	.88	1.00						
8. STYLE AND TONE	.84	.80	.73	.80	.84	.81	.82	1.00					
9. Choice and consistency of tone	.82	.78	.70	.78	.82	.80	.78	.97	1.00				
10. Choice of words and phrases	.82	.79	.72	.77	.82	.78	.81	.97	.89	1.00			
11. INTEREST IN COMPOSITION	.72	.75	.72	.71	.73	.72	.70	.70	.71	.65	1.00		
12. SENSE OF BEING PERSUADED	.66	.68	.63	.66	.67	.65	.64	.67	.68	.63	.87	1.00	
13. OVERALL RESPONSE	.71	.74	.70	.70	.73	.71	.69	.71	.72	.66	.97	.97	1.00

TABLE 4. Intercorrelations among various rating categories. (Task 7, 9 raters and 10 compositions)

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. OVERALL IMPRESSION	1.00												
2. QUALITY AND SCOPE OF CONTENT	.86	1.00											
3. Significance of what is said	.77	.94	1.00										
4. Penetration in exploration of ideas, depth	.84	.95	.78	1.00									
5. ORGANIZATION AND PRESENTATION OF IDEAS	.80	.75	.69	.72	1.00								
6. Organization of the whole text	.79	.71	.65	.69	.97	1.00							
7. Organization of sub-units	.76	.73	.68	.70	.97	.87	1.00						
8. STYLE AND TONE	.75	.72	.64	.71	.76	.72	.74	1.00					
9. Choice of words, phrases	.74	.68	.59	.68	.72	.61	.71	.94	1.00				
10. Consistency of tone	.68	.68	.62	.66	.72	.68	.71	.95	.78	1.00			
11. INTEREST IN COMPOSITION	.66	.67	.62	.64	.64	.61	.63	.65	.64	.60	1.00		
12. SENSE OF CONNECTION WITH WRITER	.66	.67	.61	.66	.64	.60	.63	.62	.58	.58	.89	1.00	
13. OVERALL RESPONSE	.68	.69	.62	.67	.66	.62	.65	.65	.62	.61	.97	.97	1.00

translated, it was decided not to include Grammatical Features, Spelling and Neatness in the rated categories. These categories will be rated according to national criteria anyway.

The results shown in Tables 1 and 2 indicate a high degree of interrater agreement (coefficient alpha, which indicates the average ratings of raters on categories, i.e. homogeneity of ratings). The overall impression was .94 in Task 6 and .93 in Task 7. Most of the coefficients in the various sub-categories are also above .90. The lowest agreement is in the rater's response to the compositions.

The pilot results are very encouraging. It seems possible to reach a high agreement between raters who come from different countries when they rate compositions written by school students in several countries. If further pilot work confirms the results in all other tasks as well, we may conclude that there tends to be a relatively high consensus among the mother tongue teaching profession of what constitutes good vs. poor writing in school-based writing.

The intercorrelations between the different rating categories exhibit a clear pattern (Tables 3 and 4). All intercorrelations are relatively high, ranging from a low of .58 to a high of .97. The overall impression has the highest correlation with the quality of ideas followed by organization, style and tone and the lowest correlation with the response of rater. The inter-correlation of each of the sub-category with its main category is very high, of the order of .94 to .97 whereas the two sub-categories typically correlate clearly lower, the range being from .78 to .89. The pattern is consistent and sensible. One would expect sub-categories to correlate highly with their overall category and less highly (but still exhibiting a considerable agreement) with each other. Since the reliabilities on each sub-category are higher than their intercorrelations it seems justified to keep the sub-categories separate rather than just have the overall category. Keeping them separate appears to add some reliable information about the compositions.

The correlations across main categories and their sub-categories are typically lower and clearly lower than the respective reliability coefficients. Thus, the several aspects of compositions are related: good content tends to go with good organization and appropriate style but the relationship is not

perfect. There can be several different profiles behind an identical overall impression grade.

7. Conclusion

It has been attempted to show how the dependent variable of the IEA International Study of Written Composition, which can in a broad sense be defined as "student writing', is conceptually divisible into two major sub-constructs: "writing competence" and writing preference."

Writing competence is defined to consist of "discourse-structuring competence" and "text-producing competence." Discourse-structuring competence requires "cognitive competence" to generate and organize ideas and "social competence" to apply appropriate style and tone in view of the readers. Text-producing competence consists of "linguistic competence" needed to use appropriate grammatical, lexical and spelling conventions and of "motor competence" needed to produce a legible text.

Since it is assumed that all tasks require from students such types of competence, it is also assumed that all tasks can be rated using a basically similar sub-set of rating categories. This does not mean, however, that all tasks are identical. Nor does it mean that the content, or meaning, of the identical rating categories is identical. The categories are the same but their meaning is specified separately task by task and may vary only slightly or quite considerably. Thus the proper form of "idea organization" is not assumed to be the same in a narrative task and in an argumentative task. The meaning and the weight of the common rating categories have to be specified task by task. The rating criteria have, as it were "structural meaning" and they need further specification to acquire also "semantic meaning."

A small scale pilot study in which mother tongue experts from nine countries rated argumentative and reflective compositions written by students from a number of countries showed that there was a high agreement on all the rating categories. This, together with the fact that the raters were able to communicate relatively easily with each other using the categories without elaborate and extended terminological discussions and definitions, suggests that the categories appear to be familiar and appropriate in various parts of the world. Thus the results lend support to the hypothesis that the conceptual

analysis of the construct "writing competence" is also empirically valid and that this validity appears to transcend cultural and linguistic boundaries. It may be that school-based writing tends to share similar characteristics and criteria irrespective of the school system. It is also possible that this similarity is higher than in non-school-based writing or in oral "texts." This hypothesis, that schools may be more alike than cultures in general and in terms of written composition than out-of-school language usage, would be an interesting hypothesis to explore further.

"Writing preference", the other aspect of "student writing", was also found to be conceptually complex. It can be argued that it can be treated both as an independent variable and thus may partly explain writing performance. On the other hand, writing preference can be considered to be, at least in part, influenced by writing instruction in school and thus it could be treated as a dependent variable in addition to writing competence. Writing preference can be measured both by means of questionnaires (self-reporting) and by analyzing students' compositions in a descriptive manner. Content analysis can then reveal what students prefer to write about, how they prefer to organize the content, what kind of style and tone they typically choose, and what kind of language they use.

In conclusion, it is argued that for teachers, teacher educators, curriculum designers, textbook writers, etc. it would be very useful if assessments of writing competence were supplemented by descriotive accounts of salient characteristics found in the compositions. To produce a conceptually sound system (with categories at several levels to reflect the complexity of the activity involved in writing a composition) is a task that deserves high priority by researchers interested in the study and assessment of writing. This paper has only attempted to argue the case for such a complementary approach without being able to make much progress in outlining the specifics of such an enterprise.

References

Brewer, W. F. and Lichtenstein, E. H. (1982) Events Schemas, Story Schemas, and Story Grammars. In J. Long and A. Baddeley (Eds.) Attention and Performance, Vol. IX. Hillsdale, NJ: Lawrence Erlbaum, 363-379.

Bruner, J. S. (1972) The Relevance of Education. Harmondsworth: Penquin Books.

Emig, J. (1977) Writing as a Mode of Learning. College Composition and Communication 28 (2), 122-128.

Purves, A. C. (1981) Rationale for the Design of the Study. IEA/WR/A2. Unpublished paper.

Scardamalia, M. and Bereiter, C. (1981) Development of Dialectical Processes in Composition. Paper presented at Conference on the Nature and Consequences of Literacy, OISE, Toronto, October 1981.

Takala, S. (1982) Developing a System for a Descriptive Account of Written Compositions. Unpublished paper.

Vähäpassi, A. (1982) On the Specification of the Domain of School Writing. Evaluation in Education 5 (3), 265-289.